



管理ツール 操作マニュアル

S01-02 データ加工設定

株式会社 NTTデータ NJK

- このソフトウェアの著作権は、株式会社 NTT データ NJK にあります。
- このソフトウェアおよびマニュアルの一部または全部を無断で使用、複製することは法律で禁止されております。
- このソフトウェアおよびマニュアルは、本製品の使用許諾契約書のもとでのみ使用することができます。
- このソフトウェアおよびマニュアルを運用した結果の影響については、一切責任を負いかねますのでご了承ください。
- このソフトウェアの仕様およびマニュアルに記載されている事項は、将来予告なしに変更することがあります。
- このマニュアルで使用している画像には、開発中のものも含まれます。実際の製品と異なる場合がございますことをご了承ください。

DataNature および DataNature Smart は、株式会社 NTT データ NJK の登録商標です。
その他の商品名は、一般に各社の商標または登録商標です。

CONTENTS

マニュアルの表記	S01-02_0
1.DataNature Smart 全体概要図	S01-02_1
2.基本操作	S01-02_2
2-1 ジョブの作成	S01-02_2
◆加工条件設定	S01-02_2
Step1: 処理対象データの指定	S01-02_2
Step2: 処理設定	S01-02_4
Step3: 出力設定	S01-02_9
◆加工条件ファイルの保存	S01-02_10
◆ジョブの設定	S01-02_11
2-2 加工条件の修正	S01-02_12

マニュアルの表記

◆ 呼称について

本マニュアルでは以下の呼称を使用しています。

DataNature Smart 管理ツール :DN 管理ツール
DataNature Smart クライアント :DN クライアント

◆ 画面に表示されるコマンド名などの文字

コマンド名やダイアログボックス名など、画面上の固有の文字は [] で囲んで表記します。なお、ユーザーによって異なるデータボタンは < > で囲んで表記します。

【例】

[編集]コマンド、[検索]ダイアログボックス
<大阪営業所>データボタン

◆ マニュアルで使用しているアイコン



“注意”を意味します。

知っておくべき重要な事柄や行なってはならない操作など、注意すべきことについて記載されています。



“ポイント”を意味します。

操作や内容のポイントとなる事柄について記載されています。



“参考”を意味します。

知っておくと便利な操作や内容などのノウハウについて記載されています。



“参照”を意味します。

操作や内容に関連するマニュアルの参照ページについて記載されています。

◆ マニュアルで使用している画像について

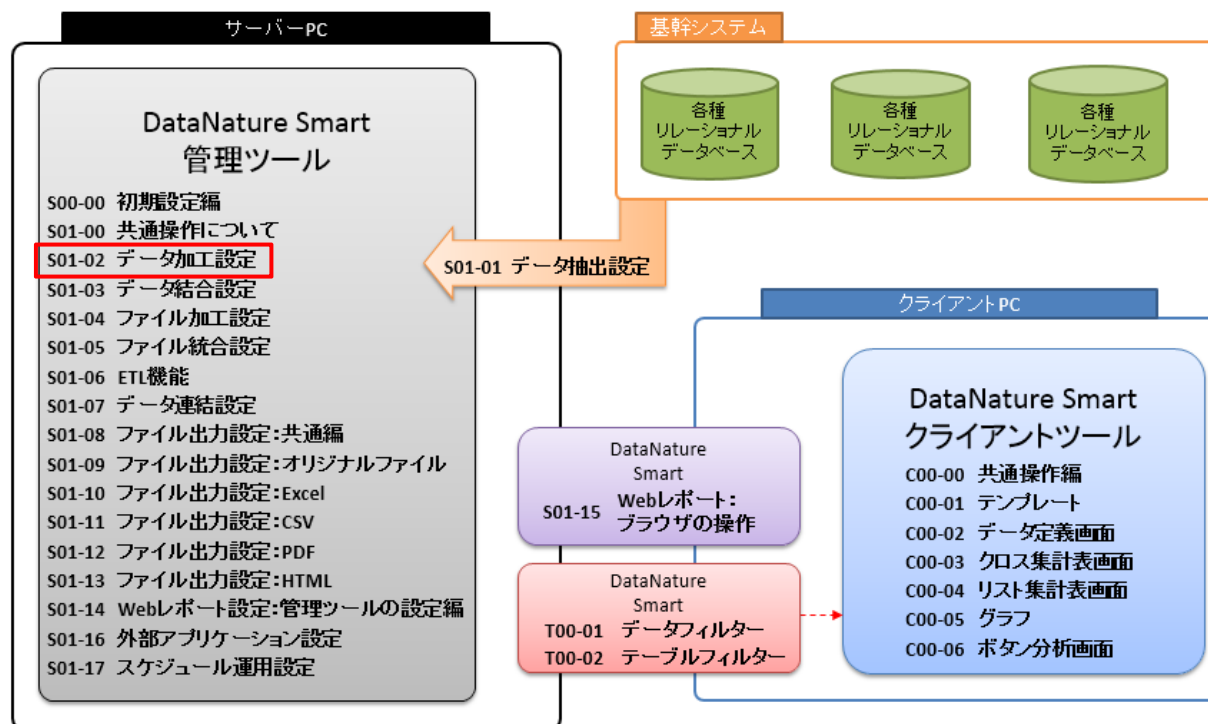
マニュアルで使用している画像には、開発中のものも含まれます。実際の製品と異なる場合がございますことをご了承ください。

◆ 機能について

導入された製品構成によって、利用できない機能や制限があります。

1. DataNature Smart 全体概要図

本マニュアルでは、S01-02 データ加工について記載しています。



CSV ファイルに対し、データの整形や新たな項目を追加するなど、データを使いやすくするための処理を設定します。

データ・クレンジング

- **無効データの排除**
空白や特定の文字列が入っているレコードを削除
- **トリミング**
スペース(空白文字)の削除
- **表記統一**
ひらがな/カタカナ、大文字/小文字、全角/半角
- **日付書式の変換**
文字列等を日付型に変換
- **データ置換**
文字列の置換(変換テーブルの作成、インポートが可能)
- **コード変換**
文字コード変換、改行コードの最適化

データ加工

- **項目間の計算**
四則演算、IF 関数、日付計算
- **項目分割**
1つの項目を、桁数、区切文字を指定し分割
- **項目結合**
2つの項目を合わせ、1つの項目に結合
- **ランクの設定**
数値にランクを設け、ランク毎の名称を付加
- **補助項目行の追加**
特定の文字列やシステム日付等を全レコードに付加
- **期間項目の設定**
年度、上期/下期、四半期、年月、月日、日、週、曜日、曜日

2. 基本操作

2-1 ジョブの作成

データ加工設定では、カンマ区切りテキストファイル(*.csv) やタブ区切りテキストファイル(*.txt)を元に、データ内容のクレンジング処理(表記統一、空白削除、無効データの排除…など)や、加工処理(計算、項目分割/結合、置換、補助項目の追加…など)を行い、分析用データを調整することができます。

◆ 加工条件設定

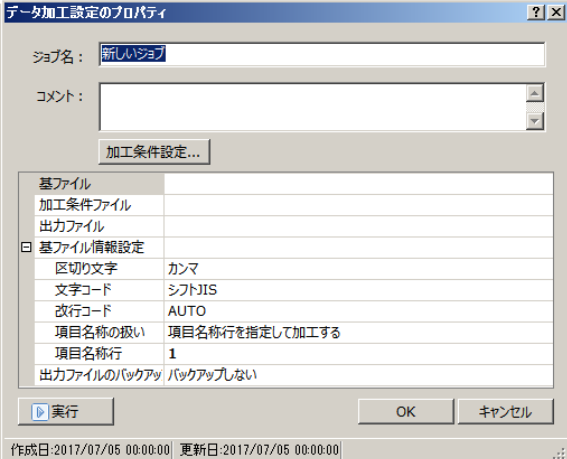
データ加工を行う基となるテキストファイル(*.csv/*.*.txt)を指定し、加工処理を設定します。

Step1: 処理対象データの指定

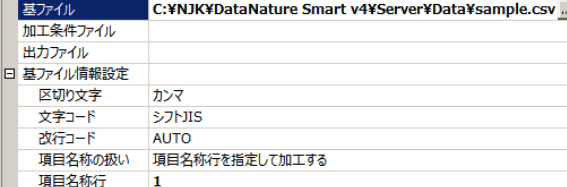
- 1  DN 管理ツールの[DataNature の管理]-[ジョブ設定]ノードをクリックし、メニューバー[ジョブ設定]-[ファイル操作系作成]-[データ加工設定]をクリックします。

([DataNature の管理]-[ジョブ設定]-[データ加工設定]ノード上のコンテキストメニュー[新しいジョブ])

▼[データ加工設定のプロパティ]が表示されます。



- 2 [ジョブ名:]テキストボックスに、ジョブ名を入力します。
- 3 [コメント:]テキストボックスに、データ加工設定についての説明などを任意に入力します。
- 4 [基ファイル]欄に、加工処理の基となるテキストファイル(*.csv/*.*.txt)を指定します。
▼[基ファイル情報設定]欄に表示されている内容とテキストファイルのフォーマット形式が異なる場合には変更してください。



- 5 [加工条件設定]ボタンをクリックします。

▼[データ加工設定]ダイアログボックスが表示されます。

- 6 基ファイルにあるデータ項目から、出力対象とする項目を選択するため、[一括設定]ボタンをクリックします。



[一括設定]ボタンは、生成列が一つもない状態でしか利用できません。

▼[データ加工設定[一括設定]]ダイアログボックスが表示されます。

- 7 [基ファイルの項目]欄に表示されている項目リストから、生成列に追加したい項目をクリックして選択し[→]ボタンで追加します。[Ctrl]キーを押しながら項目を選択すると、複数の項目を同時に選択できます。

▼選択した項目が[追加する項目]欄の項目リストへ追加されました。



[データ加工設定[一括設定]]ダイアログボックスでは、選択したすべての項目に対して、トリミングや文字変換などの表記統一の処理を一括で行うことができます。

- 8 [OK]ボタンをクリックします。
▼生成列表示領域に、追加した項目の列が表示されます。

Step2: 処理設定

データ加工設定では、以下のような処理を行うことができます。

データ加工処理の種類		処理の概要
クレンジング	データ	空白の削除、全角/半角変換といった表記統一や、無効データの処理、桁合わせを行う
	日付	日付データを日付項目として扱えるように変換処理を行う
計算式	数値計算	数値データを対象に四則演算、IF関数の計算を行う
	日付計算	2つの日付項目で減算を行う
	時間計算	2つの時間項目で減算を行う
引用・項目結合	引用	桁数や、指定した文字または文字列でデータを分割し、指定した範囲の文字列を取り出す
	項目結合	項目列と項目列、または項目列と指定した文字列を結合し、一つの項目列にまとめる
ランク設定	自動設定	任意の最小値から最大値までを、指定した数で分割し、ランク項目列を作成する
	範囲設定	任意の範囲で分割し、ランク項目列を作成する
	ファイル設定	外部ファイルを参照し、ランク項目列を作成する
補助項目		空白だけの項目列や、システム日付/任意の文字列を設定した項目列を追加する
期間設定		日付データから、四半期、上期/下期、年度などの期間項目列を作成する
データ置換	手動設定	指定した文字列を別の文字列に置き換える
	ファイル設定	外部ファイルを参照し、指定した文字列を別の文字列に置き換える

ここでは、作成例として3種類の加工処理の手順を記載しています。

Step2-1: 期間設定(列の設定修正)	S01-02-5
Step2-2: 引用・項目結合(列の設定修正)	S01-02-7
Step2-3: 計算式(列の新規作成)	S01-02-8

Step2-1: 期間設定(列の設定修正)

- ① 加工する列を選択し、[列の設定修正]ボタンをクリックします。

▼ここでは「年月」の列を選択し、「2015/01/01」の表示を「2015 年 01 月」に変換する設定を行います。

	1	2	3	4	5	6	7	8	9	10
1	年月	店舗名CD	店舗名	メーカー名	分類名	メーカー名	商品名	数量	売上金額	粗利益
2	2015/01/28...	T005	広島	A V	5.1 c h...	M電気	RX-FT53	6	67200	12480
3	2013/03/29...	T004	名古屋	A V	5.1 c h...	M電気	RX-FT53	16	179200	33280
4	2015/01/31...	T004	名古屋	A V	5.1 c h...	M電気	RX-FT53	11	123200	22880

▼[データ加工設定[列の設定修正]]ダイアログボックスが表示されます。

- ② 画面右のタブの中から、目的のデータ加工処理を選択します。ここでは、期間設定の処理を行います。



[一括設定] (S01-02_3)にて表記統一を設定した場合、[列作成の処理の順番]の[処理名称]には、「クレンジング(一括設定)」と表示されます。

- ③ [期間設定]タブをクリックし、以下のように指定します。

▼[対象列のデータ形式]領域→「日付データ」にチェックオン。

[期間の選択]領域→[期間の種類]コンボボックスから「年月」を選択。

- ④ [追加]ボタンをクリックして処理を追加します。
▼[列作成の処理の順番]の No2 に追加した処理名「期間設定」が表示されます。
[処理結果]のプレビューを確認してください。

列作成の処理の順番

No	処理名称
1	クレンジング(一括設定)
2	期間設定

処理の削除

2015年01月
2013年03月
2015年01月
2012年12月

- ⑤ [データ加工設定[列の設定修正]]ダイアログボックスの[OK]ボタンをクリックして、設定を完了します。
▼[データ加工設定]ダイアログボックスに加工した内容が反映されます。

DataNature Smart データ加工設定

ファイル(F) 表示(V)

処理設定

列の新規作成... 列の設定修正... 列の削除 一括設定...

各ボタンにマウスを移動すると説明を表示します。

	1	2	3	4	5	6	7	8	9	10
1	年月	店舗名CD	店舗名	コーナー名	分類名	メーカー名	商品名	数量	売上金額	粗利益
2	2015年01月	T005	広島	A V	5.1 c h...	M電気	RX-FT53	6	67200	12480
3	2013年03月	T004	名古屋	A V	5.1 c h...	M電気	RX-FT53	16	179200	33280
4	2015年01月	T004	名古屋	A V	5.1 c h...	M電気	RX-FT53	11	123200	22880

Step2-2: 引用・項目結合(列の設定修正)

- ① 加工する列を選択し、[列の設定修正]ボタンをクリックします。
▼ここでは<店舗コード>と<店舗名>の列を結合し、1つの項目にまとめます。<店舗名>を選択してください。

DataNature Smart データ加工設定

ファイル(F) 表示(V)

処理設定

列の新規作成... 列の設定修正... 列の削除 一括設定...

各ボタンにマウスを移動すると説明を表示します。

	1	2	3	4	5	6	7	8	9	10
1	年月	店舗名CD	店舗名	コーナー名	分類名	メーカー名	商品名	数量	売上金額	粗利益
2	2015年01月	T005	広島	A V	5.1 c h...	M電気	RX-FT53	6	67200	12480
3	2013年03月	T004	名古屋	A V	5.1 c h...	M電気	RX-FT53	16	179200	33280
4	2015年01月	T004	名古屋	A V	5.1 c h...	M電気	RX-FT53	11	123200	22880

- ② [引用・項目結合]タブをクリックし、更に[項目結合]タブをクリックします。

クレンジング | 計算式 | 引用・項目結合 | ランク設定 | 補助項目 | 期間設定 | データ置換

引用 項目結合

結合内容

結合内容(項目名称表示)

結合する項目の指定

表示	項目名称	種別
SD1	年月	生成列
SD2	店舗名CD	生成列
RC1	年月	CSV
RC2	店舗名CD	CSV
RC3	店舗名	CSV
RC4	コーナー名	CSV
RC5	分類名	CSV
RC6	メーカー名	CSV
RC7	商品名	CSV

+

-

入力

- ③ テキストファイルの項目を基に＜店舗コード＞と＜店舗名＞の結合を設定します。
 - ④ [結合する項目の指定]欄に表示されている項目リストから、＜店舗コード＞(種別: csv)を選択し、[入力]ボタンをクリックします。[結合内容]欄に「RC2」と表示されます。
 - ⑤ 次に、[+]ボタンをクリックします。
 - ⑥ 続いて、結合後の文字列を分かりやすく表示するために、項目の間にスペースを入力します。文字列を入力する場合は、["]ボタンで文字列を括って指定します。ここでは、["][半角スペース]["][+]と入力します。
 - ⑦ 最後に、[結合する項目の指定]欄に表示されている項目リストから、＜店舗名＞(種別: csv)を選択し、[入力]ボタンをクリックします。[結合内容]欄に「RC3」と表示されます。
- ▼[結合内容]欄には、次のように表示されます。

- ⑧ [追加]ボタンをクリックして、処理を追加します。
- ▼ [列作成の処理の順番]の No1 に「項目結合」と表示されます。[処理結果]のプレビューを確認してください。

- ⑨ [データ加工設定[列の設定修正]]ダイアログボックスの[OK]ボタンをクリックして、設定を完了します。
- ▼[データ加工設定]ダイアログボックスの＜店舗名＞には、「店舗コード(スペース)店舗名」と表示されます。

Step2-3: 計算式 (列の新規作成)

- ① 処理を新たに追加するため、[列の新規作成]ボタンをクリックします。
- ▼ここでは、計算設定を利用し、＜売上金額＞と＜粗利益＞から＜仕入金額＞の列を新規に作成します。

▼[データ加工設定[列の新規作成]]ダイアログボックスが表示されます。

- ② 項目名称を設定します。ここでは「仕入金額」と直接キーボードから入力してください。
- ③ 次に、[計算式]タブをクリックし、[数値計算]タブを表示します。

データ加工設定[列の新規作成]

項目名称
仕入金額

列作成の処理の順番

No	処理名称

処理の削除

プレビュークリア

クレンジング 計算式 引用・項目結合 ランク設定 補助項目 期間設定 データ置換

数値計算 日付計算 時間計算

計算式
計算式(項目名称表示)

項目列入力

表示	項目名称	種別
SD1	年月	生成列
SD2	店舗名CD	生成列
SD3	店舗名	生成列
SD4	コーナー名	生成列
SD5	分類名	生成列
SD6	メーカー名	生成列
SD7	商品名	生成列
SD8	数量	生成列
SD9	売上金額	生成列

計算できない時
☐ NULLとする
☐ 0とする

入力

IF関数.. () + - * /

クリア 表示

・IF関数内で文字列を使用する場合は"文字列"のように指定します。
 ・"r"を文字として使用する場合は重なる"r""と指定します。

クリア 修正 追加

OK キャンセル

- ④ [項目列入力]欄に表示されている項目リストから、＜売上金額＞(種別:csv)を選択し、[入力]ボタンをクリックします。[計算式]欄に「RC9」と表示されます。
 - ⑤ 次に、[-]ボタンをクリックします。
 - ⑥ 続いて、[項目列入力]欄に表示されている項目リストから、＜粗利益＞(種別:csv)を選択し、[入力]ボタンをクリックします。[計算式]欄に「RC10」と表示されます。
- ▼[計算式]欄には、次のように表示されます。

計算式
RC9-RC10

計算式(項目名称表示)
売上金額-粗利益

クリア 表示

[表示]をクリックすると、項目名称で確認できます。

- ⑦ [追加]ボタンをクリックして、処理を追加します。
- ▼[列作成の処理の順番]の No1 に「計算式(数値計算)」と表示されます。[処理結果]のプレビューを確認してください。

列作成の処理の順番

No	処理名称
1	計算式(数値計算)

処理の削除

54720
145920
100320

- ⑧ [データ加工設定[列の設定修正]]ダイアログボックスの[OK]ボタンをクリックして、設定を完了します。
- ▼[データ加工設定]ダイアログボックスの生成列の最後尾に＜仕入金額＞が表示されます。

DataNature Smart データ加工設定

ファイル(F) 表示(V)

処理設定
列の新規作成... 列の設定修正... 列の削除 一括設定...

各ボタンにマウスを移動すると説明を表示します。

	3	4	5	6	7	8	9	10	11	12
1	店舗名	コーナー名	分類名	メーカー名	商品名	数量	売上金額	粗利益	予算額	仕入金額
2	F005 広島	A V	5.1 c h ...	M電気	RX-FT53	6	67200	12480		54720
3	F004 名古屋	A V	5.1 c h ...	M電気	RX-FT53	16	179200	33280		145920
4	F004 名古屋	A V	5.1 c h ...	M電気	RX-FT53	11	123200	22880		100320

Step3: 出力設定

加工処理の設定が完了したら、テキストファイル(*.csv)に出力する項目を設定します。

- 1 [データ加工設定]ダイアログボックス下段の[出力設定]領域にある[出力列の選択]ボタンをクリックします。
▼[データ加工設定[出力列の選択]]ダイアログボックスが表示されます。

- 2 [生成した項目]欄に表示されている項目リストから、出力したい項目をクリックして選択し、[→]ボタンで追加します。[Cntl]キーを押しながら項目を選択すると、複数の項目を同時に選択できます。
- 3 Step 2-2 で項目結合を行い、列 No3 の<店舗名>の内容を<店舗コード+店舗名>としているので、列 No2 の店舗コードは不要です。[出力する項目]欄から<店舗コード>を選択して[←]ボタンでクリアします。



右端の[▲]や[▼]ボタンを利用して、出力する項目の並び順を変更することができます。

- 4 [出力プレビュー表示]ボタンをクリックします。
▼クリアした<店舗コード>の項目は表示されていません。

No.	1	2	3	4	5	6	7
1	年月	店舗名	店舗名	店舗名	店舗名	店舗名	店舗名
2	2015年01月	T005 広島	A V	5.1 c hスピーカー	M電気	RX-FT53	6
3	2013年03月	T004 名古屋	A V	5.1 c hスピーカー	M電気	RX-FT53	16
4	2015年01月	T004 名古屋	A V	5.1 c hスピーカー	M電気	RX-FT53	11

- 5 プレビュー画面を閉じて、[データ加工設定[出力列の選択]]ダイアログボックスの[OK]ボタンをクリックします。
▼[データ加工設定]ダイアログボックスに戻ります。出力する項目から外した2列目の<店舗コード>はグレイ表示になっています。

No.	1	2	3	4	5	6	7	8	9	10
1	年月	店舗名CD	店舗名	店舗名	店舗名	店舗名	店舗名	店舗名	店舗名	店舗名
2	2015年01月	T005	T005 広島	A V	5.1 c h...	M電気	RX-FT53	6	67200	12480
3	2013年03月	T004	T004 名古屋	A V	5.1 c h...	M電気	RX-FT53	16	179200	33280
4	2015年01月	T004	T004 名古屋	A V	5.1 c h...	M電気	RX-FT53	11	123200	22880

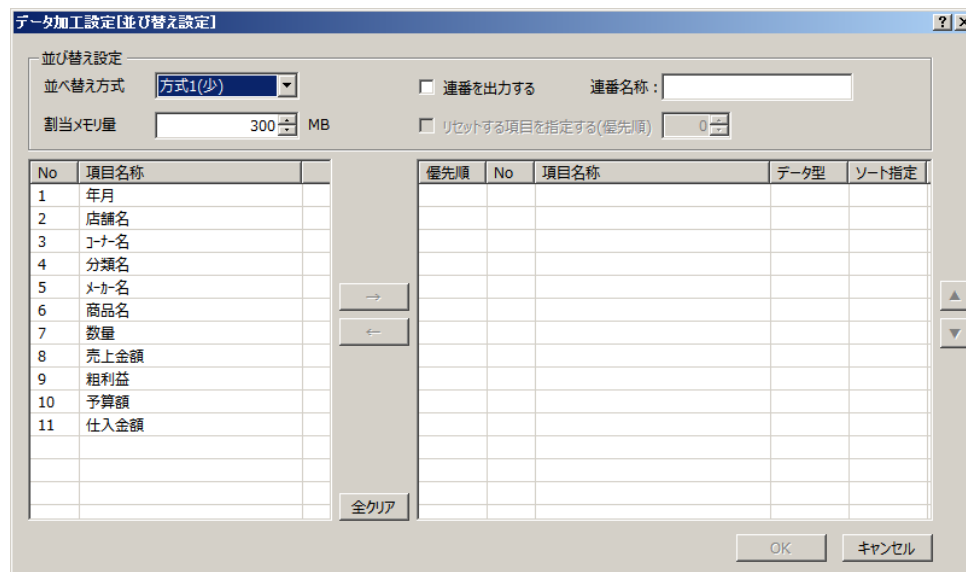
**並び替え設定について**

データ加工では、出力設定の際、レコードの並び替えと連番の出力設定が行えます。

設定を行なう場合は、「出力設定」欄の[並び替え設定]をチェックオンし、[並び替え設定]ボタンをクリックします。

No.	項目名称	優先順	データ型	ソート指定
1	年月			
2	店舗名			

[データ加工設定[並び替え設定]]ダイアログボックスが表示され、並び替えに関する設定が行えます。



「並び替え設定」欄では、並び替え方式やこの処理で使用するメモリの最大量、連番の出力有無（最終列に新たな項目を追加し、並び替えた結果の順番を出力）等を指定します。

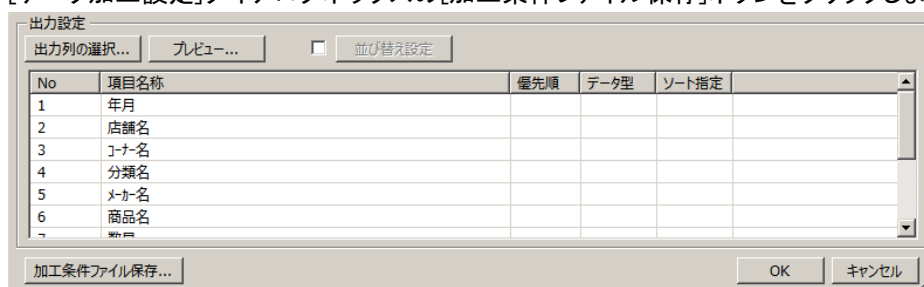
並び替え対象の項目については、項目一覧から選択し[→]ボタンで指定します。優先順や連番を出力する際のデータ型（文字/数値）、ソート指定（昇順/降順）が行えます。

設定ダイアログボックスの詳細については、ヘルプを参照してください。

◆ 加工条件ファイルの保存

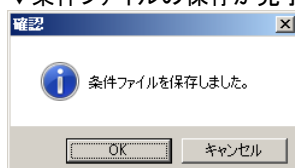
設定した加工処理を、データ加工条件ファイル(*.clex)に保存します。

- 1 [データ加工設定]ダイアログボックスの[加工条件ファイル保存]ボタンをクリックします。



- 2 [名前を付けて保存]ダイアログボックスで、保存先を指定しファイル名を入力して、[保存]ボタンをクリックします。

▼条件ファイルの保存が完了しました。[OK]ボタンをクリックします。



- ③ [データ加工設定]ダイアログボックスの[OK]ボタンをクリックします。
▼[データ加工設定のプロパティ]ダイアログボックスに戻ると、[加工条件ファイル]欄には、今保存したファイル名が表示されています。

◆ ジョブの設定

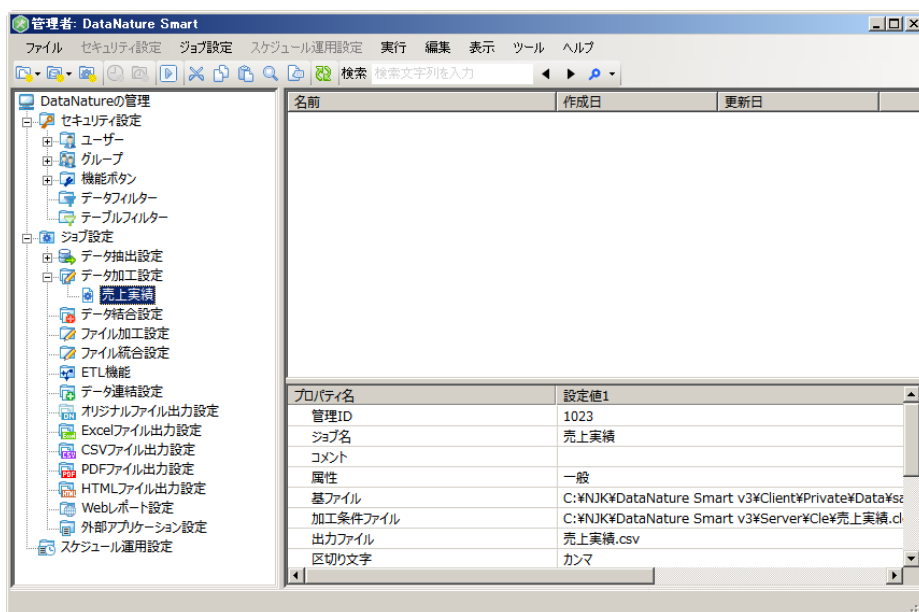
データ加工条件ファイル(*.clex)を利用して、ジョブの設定を行います。

- ① [データ加工設定のプロパティ]ダイアログボックスで、処理対象となるファイルを指定します。基ファイルと加工条件ファイルは既に指定されているので、出力ファイルを設定してください。
▼ジョブの全ての情報が設定されました。



[実行]ボタンをクリックすると、[読み込み件数の設定]ダイアログボックスが表示され、件数を限定してテスト実行することができます。

- ② 設定が完了しましたので、[OK]ボタンをクリックしてジョブを登録します。
▼[データ加工設定のプロパティ]ダイアログボックスが閉じ、管理ツールの画面に戻ります。設定した内容がプロパティ表示領域に表示されます。

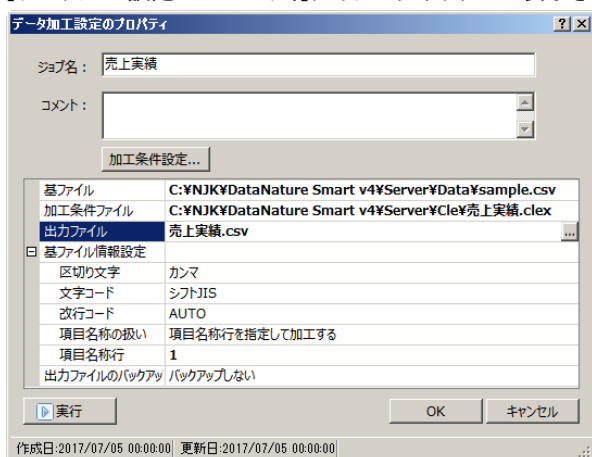


2-2 加工条件の修正

保存済みのデータ加工条件ファイル(*.cle、*.clex)の修正を行います。

※旧 DataNature 製品で作成した cle 形式の加工条件ファイルの読み込みは可能ですが、出力は clex 形式のみとなります。

- 1 DN 管理ツールの[DataNature の管理]-[ジョブ設定]-[データ加工設定]ノード上に登録されている処理の中から修正を行いたいジョブ上で右クリックし、[プロパティ]をクリックします。
[データ加工設定のプロパティ]ダイアログボックスが表示されます。



- 2 [加工条件設定]ボタンをクリックします。
▼[データ加工設定[***.clex]]ダイアログボックスが表示されます。



加工処理の内容を確認するには、生成列表示領域から対象の列を選択し、[列の設定修正]ボタンをクリックします。[列作成の処理の順番]の各処理を選択すると、設定内容が画面右側に表示されます。



[基ファイル]に指定したデータファイルのプレビューを表示し、データ加工の基になる項目の内容を確認することができます。メニューバー[表示]-[基ファイルプレビュー表示]から実行してください。[データ加工設定 [基ファイルプレビュー][基ファイル名]]ウィンドウが表示され、基ファイルに指定したデータファイルの上位 100 件のプレビューが確認できます。

- ③ 任意に加工条件の修正を行います。修正方法は、Step 2: 処理設定 (S01-02_4) や各加工処理のヘルプを参照してください。
- ④ 編集が完了したら、加工条件ファイル(*.clex)を保存してください。